WO 98/20120 PCT/US97/20560

Fig. 1A

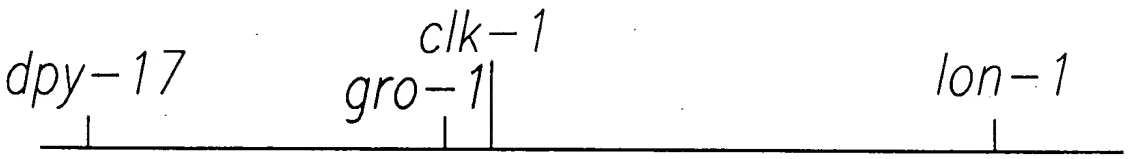
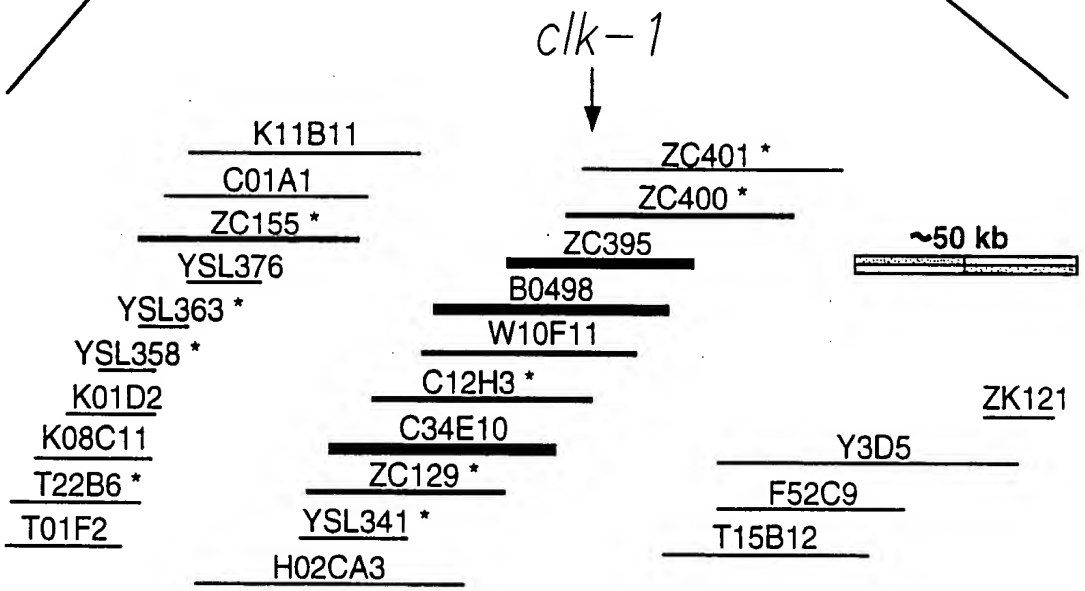



Fig. 1B



005220 "TSTCT560

Rescue

Fig 2A

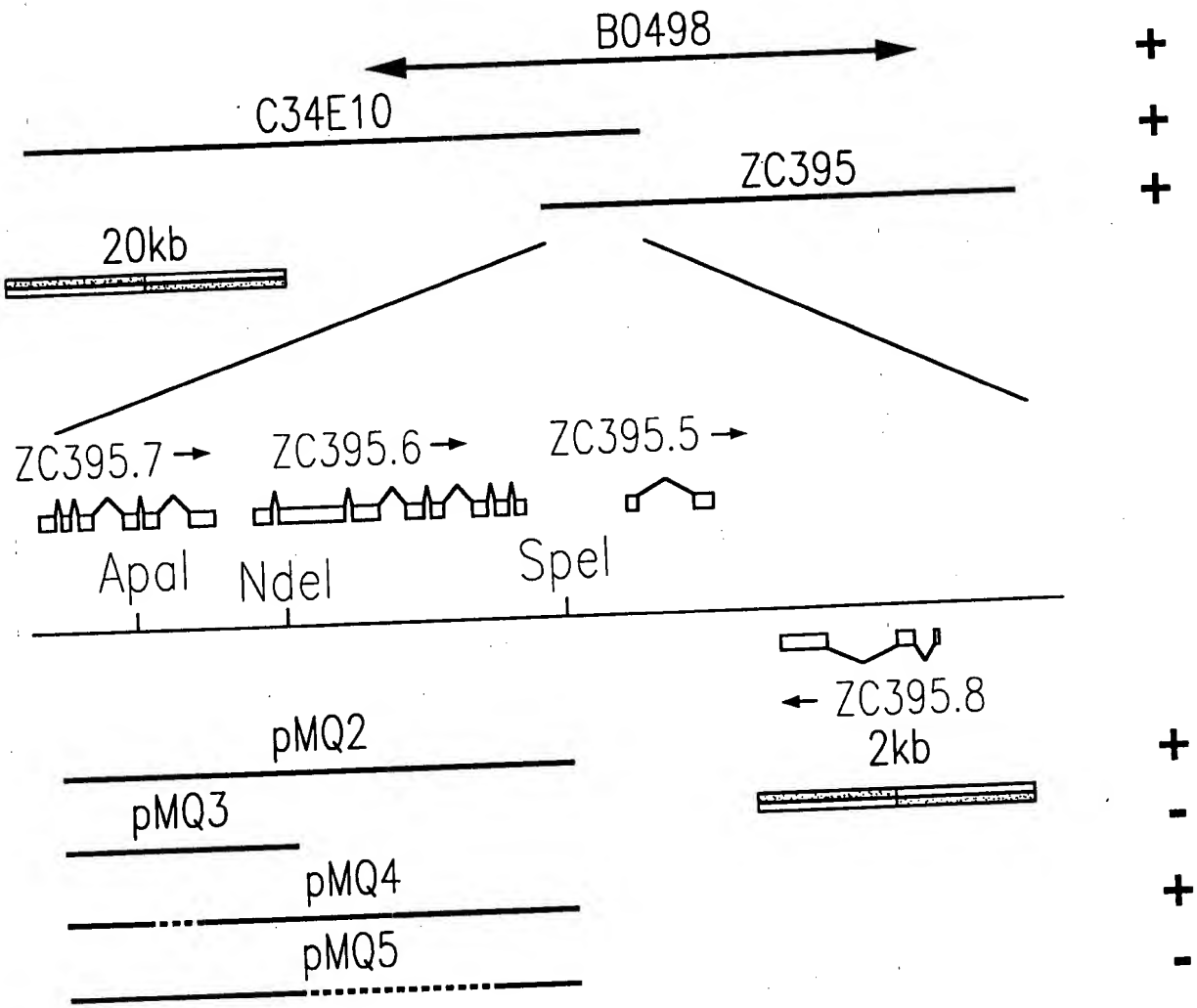



Fig 2B

*gro-1*

SL2

M I F R K F L N F L K P Y K M R 16

aaaatatcgtcaggaaataataacatttcagatataccctgaactctacagtttATGATATTCAGGAAATTTCTGAATTTTCTGAAACCTTACAAAATGC 1394

T D P I I F V I G C T G T G K S D L G V A I A K K Y G G E V I S V 49

GAACGGATCCGATTATTTTCGTGATTGGGTGCACTGGAACCGGGAAAAGTGATCTTGGAGTGGCAATTGCAAAGAAATATGGAGGAGAGGTGATTAGTGT 1494

SHP109

D S M Q F Y K G L D I A T N K I T 66

AGATTCAATGCAATTTTATAAAGgtacatgggttttgtttcaattttaaattaattaattttcgtttttcagGACTTGACATTGCCACGAATAAGATAAC 1594

E E E S E G I Q H H M M S F L N P S E S S S Y N V H S F R E V T L 99

GGAAGAAGAATCTGAAGGGATTCAACATCATATGATGTCATTTTTGAATCCATCTGAATCATCATCTTATAATGTACATAGTTTCCGAGAAGTCACGTTG 1694

SHP94

D L I K K I R A R S K I P V I V G 116

GATCTTATTAAAgtgcttaattcgccacttttttgaacttgatcctaattttcataattttcagAAAATCCGCGCCCGTTCAAAAATTCCTGTAATTGTCG 1794

SHP95

G T T Y Y A E S V L Y E N N L I E T N T S D D V D S K S R T S S E 149

GAGGAACCACTTATTATGCTGAAAGTGTCCTTTATGAGAATAATCTGATTGAAACCAACACTTCAGATGACGTGGATTCCAAATCGAGAACATCATCAGA 1894

SHP96

S S S E D T E E G I S N Q E L W D E L K K I D E K S A L L L H P N 182

ATCGTCATCTGAAGACACTGAAGAAGGAATTAGTAATCAAGAATTATGGGATGAATTGAAAAAAATCGACGAAAAATCAGCACTTCTTCTACATCCAAAT 1994

FIG 3A

## *gro-1* continued...

```
N R Y R V Q R A L Q I F R E T G                                                  198

AATCGTTATCGAGTACAGAGAGCATTGCAAATTTTCAGAGAAACTGgtaattgatttgcaaatttccagattaaaaacaaatcaagtaaagttttttgca  2094

  I R K S E L V E K Q K S D E T V D L G G R L R F D N S L V I F M D  231

gGAATCCGAAAAAGTGAACTTGTTGAAAAACAGAAATCAGATGAAACTGTTGATTTGGGTGGACGACTACGATTTGATAATTCTTTAGTTATTTTTATGG  2194
SHP97

  A T P E V L E E R L D G R V D K M I K L G L K N E L I E F Y N E    263

ATGCAACACCTGAAGTTTTAGAAGAAAGACTTGATGGAAGAGTTGATAAAATGATTAAATTGGGTTTGAAGAATGAATTGATCGAGTTTTATAACGAGgt  2294

aaatatttgaatttttccagaaaaaaaaagaaaatttttattattttgttttttttttcattctttactattttccaaaaaagtttaaacttttgaaaac  2394

                                                                  H A E Y  267

tgttcagaaaatgttcgtgtatttatttttagcttactgaggcattatttcattgtgatttttactatactctataaactaaatttttcagCACGCCGAGTA  2494

I N H S K Y G V M Q C I G L K E F V P W L N L D P S E R D T L N G  300

CATAAATCACAGCAAATATGGTGTCATGCAATGTATTGGTCTTAAAGAATTCGTTCCATGGCTCAATTTGGACCCATCAGAAAGAGATACACTCAATGGG  2594
                          CG  e2400 lesion                                  SHP98

D K L F K Q G                                                  C D D V K L H T R Q Y  318

GATAAATTGTTCAAGCAAGGgtaatttaaatttattttcaatttttataaattccaagctattttcagATGCGATGATGTGAAGCTTCACACTCGACAAT  2694
```

Fig 3B

*gro-1* continued...

```
 A  R  R  Q  R  R  W  Y  R  S  R  L  L  K  R  S  D  G  D  R                    33

ATGCACGGCGCCAGAGACGGTGGTATCGATCGAGACTTTTAAAACGGTCGGATGGTGATCGGgtatgttgattttaaaaaaattgaatttttaaagaact  279
          SHP99

tttttactaaattaacaaagttattggctgaaaatggctgaaaattatagtaaaactaatcaaaaaaattgaaattttgaattaaagtcataaagtgacg  289

                                                                 K  M  A  S  T  K  M  L  D   34

accagaaaattaaaaaaaaacatttttctattttaattaattcactctacttcactttaaaaataattttcagAAAATGGCAAGTACAAAAATGCTGGAT  299

 T  S  D  K  Y  R  I  I  S  D  G  M  D  I  V  D  Q  W  M  N  G  I  D  L  F  E  D                    37

ACATCTGACAAGTACCGAATAATTAGTGATGGAATGGACATTGTTGATCAATGGATGAATGGAATCGATCTATTTGAAGATgtaaaatttcacaaattct  309

                                                 I  S  T  D  T  N  P  I  L  K  G  S  D  A  N  I  L  L  N  C  E  I   39

aaaatttccgaatcacaaattaaaatttctacagATCTCCACAGACACCAATCCAATTCTAAAAGGGTCCGATGCAAATATTCTGCTGAATTGTGAAATC  319

 C  N  I  S  M  T  G  K  D  N  W                                    Q  K  E  I  D  G  K  K   41

TGTAATATTTCAATGACTGGAAAAGATAATTGgtttgtttcaatacatattataatttcgaaatgaatttttttcagGCAGAAACATATCGATGGGAAAAA  329
SHP110      SHP100

 H  K  H  H  A  K  Q  K  K  L  A  E  T  R  T  •                                                      43

GCACAAGCATCATGCTAAGCAAAAGAAATTGGCAGAGACTCGCACAtaagacgctatatttatttttttgttaacttaaattatttttgttgttgattgtt  339

                    polyA

ctctaaataaaaaaacagctcagagagaagattaggcgctcgtccacatctccgacgatagtcaacccgaacgaagggaactatctttaattgtcagtga  349
                                                                                SHP92
```

FIG. 3C

```
tgatttttactatactctataaactaaattttcagCACGCCGAGTACATAAATCACAGCAAATATGGTGTCACG      1197
                                  H  A  E  Y  I  N  H  S  K  Y  G  V  T        276

   TTGGTCTTAAAGAATTCGTTCCATGGCTCAATTTGGACCCATCAGAAAGAGATACACTCAATGGGGATAAATTGT  1272
   L  V  L  K  N  S  F  H  G  S  I  W  T  H  Q  K  W  I  H  S  M  G  I  N  C    301

TCAAGCAAGGgtaatttaaatttattttcaatttttataaattccaagctattttcagATGCGATGATGtgaagcttc  1350
S  S  K  D                                                 A  M  M  •        308
```

Fig 30

SUBSTITUTE SHEET (RULE 26)

005220 "PCT/US60

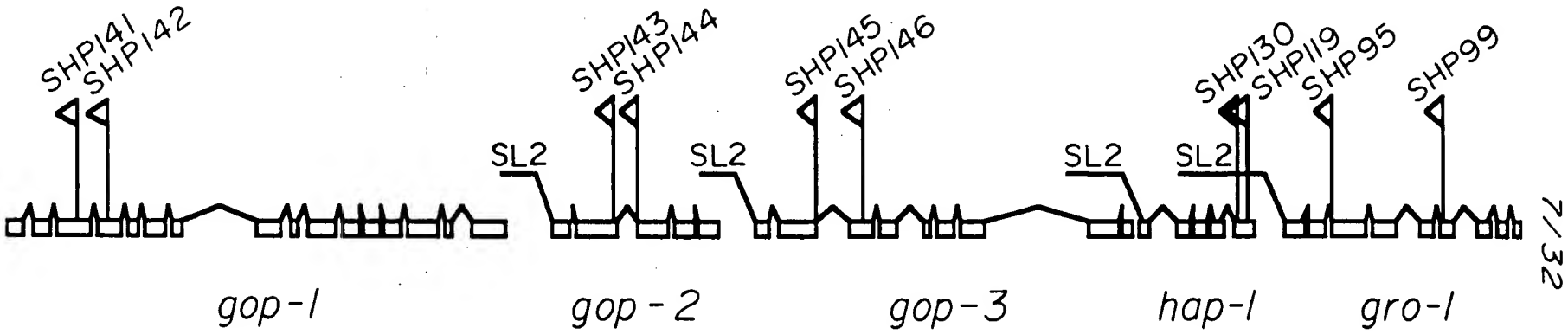





FIG-4A

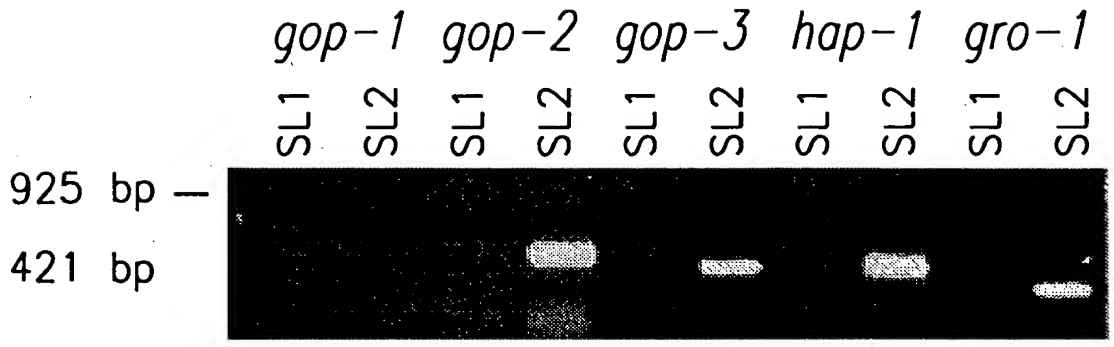

gop-1
gop-2
gop-3
hap-1
gro-1
SL1
SL2
SL1
SL2
SL1
SL2
SL1
SL2
SL1
SL2
925 bp
421 bp

Fig. 4B

# Sequence of GRO-1 and homologues

```
                                 .  . .  ..  .        .   ..  .   . . .   .
C.elegans        1 MIFRKFLNFLKPYKMRTDPIIFVIGCTGTGKSDLGVAIAKKYGGEVISVDSMQFYKGLDIATNKITEEESEGIQ
S.cerevisiae     1    MLKGPLKGCLNMSKKVIVIAGTTGVGKSQLSIQLAQKFNGEVINSDSMQVYKDIPIITNKHPLQEREGIP
E.coli           1        MSDISKASLPKAIFLMGPTASGKTALAIELRKILPVELISVDSALIYKGMDIGTAKPNAEELLAAP

                                    ATP/GTP
                                    binding site

                    .   .   .        .    ..  ....  .
C.elegans       76 HMMSFLNPSESSSYNVHSFREVTLDLIKKIRARSKIPVIVGGTTYYAESVLYENNLIETNTSDDVDSKSRTSSE
S.cerevisiae    72 HVMNHVDWSE--EYYSHRFETECMNAIEDIHRRGKIPIVVGGTHYYLQTLFNKRVDTKSSERKLTRKQLDILES
E.coli          68 RLLDIRDPSQ--AYSAADFRRDALAEMADITAAGRIPLLVGGTMLYFKALLEGLSPLPSADPEVRARIEQQAAE

                                .   .   .   ...   .  . .     .   .           .
C.elegans      151 SSEDTEEGISNQELWDELKKIDEKSALLLHPNNRYRVQRALQIFRETGIRKSELVEKQKSDETVDLGGRLRFDN
S.cerevisiae   147 DPDV---------IYNTLVKCDPDIATKYHPNDYRRVQRMLEIYYKTGKKPSETFNEQK--------ITLKFD-
E.coli         143 GWES---------LHRQLQEVDPVAAARIHPNDPQRLSRALEVFFISGKTLTELTQTSG--------DALPYQV
```

FIG 5A

e2400

```
                                                                    |
                                  .  .  .  .  .  .  .               |
C.elegans        226 LVIFMDATPEVLEERLDGRVDKMIKLGLKNELIEFYNEHAEYINHSKYGVMQCIGLKEFVPWLNLDPSERDTLN
S.cerevisiae     205 LFLWLYSKPEPLFQRLDDRVDDMLERGALQEIKQLYEYYSQNKFTPEQCENGVWQVIGFKEFLPWLTGKTDDNT
E.coli           202 QFAIAPASRELLHQRIEQRFHQMLASGFEAEVRALFARGDLHTDLPSIRCVGYRQMWSYLEGEISYDEMVYRGV

                                  ...  .  ..  .                                          .
C.elegans        301 DKLFKQGCDDVKLHTRQYARRQRRWYRSRLLKRSDGDRKMASTKMLDTSDKYRIISDGMDIVDQWMNGIDLFED
S.cerevisiae     280 KLEDCIERMKT--RTRQYAKRQVKWIKKMLIPDIKGDILLDATDLSQWDTNASQRAIAISNDFISNRPIKQERA
E.coli           277 ------------ATRQLAKRQITWLRGWEGVHWLDSEKPEQARDEVLQVVGAIAG

                                  ..     .  C2H2 zinc finger  .
C.elegans        376 STDTNPILKGSDANILLNCEICNISMTGKDNWQKHIDGKKHKHHAKQKKLATRT
S.cerevisiae     353 KALEELLSKGETTMKKLDDWTHYTRNVCRNADGKNVVAIGEKYWKIHLGSRRHKSNLKRNTRQADFEKWKINKK
```

FIG. 5B

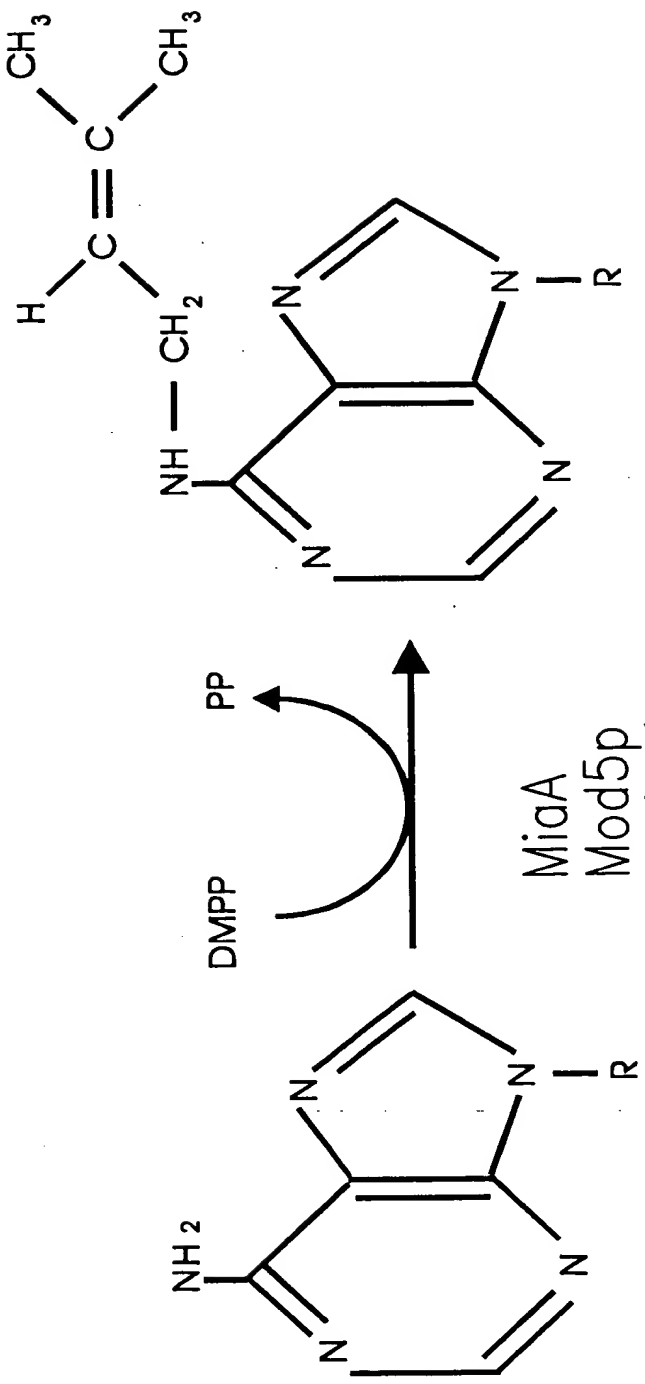



FIG-6

# Sequence of HAP-1 and its homologues

```
                                    ... . . .                                .

H. sapiens                    MAASLVGKKIVFVTGNAKKLEEVVQILGDKFP-----CTLVAQKIDLPEYXG-EPDEISIQKCQE
C. elegans          MLYILWKLNYLQKKMSLRKINFVTGNVKKLEEVKAILKNFE--------VSNVDVDLDEFQG-EPEFIAERKCRE
S. cerevisiae                 MSNNEIVFVTGNANKLKEVQSILTQEVDNNNKTIHLINEALDLEELQDTDLNAIALAKGKQ
E. coli                         MQKVVLATGNVGKVRELASLLSDFGLD-----IVAQTDLGVDSAEETGLTFIENAILKA

                                .   . .      .. .              .          .

H. sapiens          AVRQV-QG-PVLVEDTCLCFNALGKLPGPYIKWFL--EKLKPEGLHQLLAGFED-----KSAYALCTFALSTGDP
C. elegans          AVEAV-KG-PVLVEDTSLCFNAMGGLPGPYIKWFL--KNLKPEGLHNMLAGFSD-----KTAYAQCIFAYTEG-L
S. cerevisiae       AVAALGKGKPVFVEDTALRFDEFNGLPGAYIKWFL--KSMGLEKIVKMLEPFEN-----KNAEAVTTICFADSRG
E. coli             RHAAKVTALPAIADDSGLAVDVLGGAPGIYSARYSGEDATDQKNLQKLLETMKDVPDDQRQARFHCVLVYLRHAE

                                .  . .    .  .. . .      . .    .  ... .

H. sapiens          SQPVRLFRGRTSGRIV-APRGCQDFGWDPCFQP-DGYEQTYAEMPKAEKNAVSHRFRALLELQEYFGSLAA
C. elegans          GKPIHVFAGKCPGQIV-APRGDTAFGWDPCFQP-DGFKETFGEMDKDVKNEISHRAKALELLKEYFQNN
S. cerevisiae       E--YHFFQGITRGKIV-PSRGPTTFGWDSIFEPFDSHGLTYAEMSKDAKNAISHRGKAFAQFKEYLYQNDF
E. coli             DPTPLVCHGSWPGVITREPAGTGGFGYDPIFFV-PSEGKTAAELTREEKSAISHRGQALKLLLDALRNG
```

Fig. 7

**mRNA sequence of human homologue of *gro-1*: hgro-1**

CTGCCATAAG **ATG**GCGTCCG TGGCGGCTGC ACGAGCAGTT CCTGTGGGCA
GTGGGCTCAG GGGCCTGCAA CGGACCCTAC CTCTTGTAGT GATTCTCGGG
GCCACGGGCA CCGGCAAATC CACGCTGGCG TTGCAGCTAG GCCAGCGGCT
CGGCGGTGAG ATCGTCAGCG CTGACTCCAT GCAGGTCTAT GAAGGCCTAG
ACATCATCAC CAACAAGGTT TCTGCCCAAG AGCAGAGAAT CTGCCGGCAC
CACATGATCA GCTTTGTGGA TCCTCTTGTG ACCAATTACA CAGTGGTGGA
CTTCAGAAAT AGAGCAACTG CTCTGATTGA AGATATATTT GCCCGAGACA
AAATTCCTAT TGTTGTGGGA GGAACCAATT ATTACATTGA ATCTCTGCTC
TGGAAAGTTC TTGTCAATAC CAAGCCCCAG GAGATGGGCA CTGAGAAAGT
GATTGACCGA AAAGTGGAGC TTGAAAAGGA GGATGGTCTT GTACTTCACA
AACGCCTAAG CCAGGTGGAC CCAGAAATGG CTGCCAAGCT GCATCCACAT
GACAAACGCA AAGTGGCCAG GAGCTTGCAA GTTTTTGAAG AAACAGGAAT
CTCTCATAGT GAATTTCTCC ATCGTCAACA TACGGAAGAA GGTGGTGGTC
CCCTTGGAGG TCCTCTGAAG TTCTCTAACC CTTGCATCCT TTGGCTTCAT
GCTGACCAGG CAGTTCTAGA TGAGCGCTTG GATAAGAGGG TGGATGACAT
GCTTGCTGCT GGGCTCTTGG AGGAACTAAG AGATTTTCAC AGACGCTATA
ATCAGAAGAA TGTTTCGGAA AATAGCCAGG ACTATCAACA TGGTATCTTC
CAATCAATTG GCTTCAAGGA ATTTCACGAG TACCTGATCA CTGAGGGAAA
ATGCACACTG GAGACTAGTA ACCAGCTTCT AAAGAAAGGA CCTGGTCCCA
TTGTCCCCCC TGTCTATGGC TTAGAGGTAT CTGATGTCTC GAAGTGGGAG
GAGTCTGTTC TTGAACCTGC TCTTGAAATC GTGCAAAGTT TCATCCAGGG
CCACAAGCCT ACAGCCACTC CAATAAAGAT GCCATACAAT GAAGCTGAGA
ACAAGAGAAG TTATCACCTG TGTGACCTCT GTGATCGAAT CATCATTGGG
GATCGCGAAT GGGCAGCGCA CATAAAATCC AAATCCCACT TGAACCAACT
GAAGAAAAGA AGAAGATTGG ACTCAGATGC TGTCAACACC ATAGAAAGTC
AGAGTGTTTC CCCAGACTAT AACAAAGAAC CTAAAGGGAA GGGATCCCCA
GGGCAGAATG ATCAAGAGCT GAAATGCAGC GTTTAAGAGA CATGTCCAGT
GGCCTTTGGA AAGGTGGTGG GGATCCAGTT CAGGAGGGAG GGGTATGTTT
GTCTCCCAGT CTGGGCAAAG GAGTGCTATG CGGAATTCTC TGCATAGCAG
AAAAGCTCCC ACCATTTTCT TTTGATGTGG TTTTAAAGTC TCACGTTCTC
TATAATAGAA ACAGCAGGTC TTGTCAGCTC CTTGTGTGGC TGATGTGTCT
GGAAATGATG TAGTTCAGGA AAGCATTTTT TTTTTCTTTG AACCTTAAAG
GTTCTATTAT TAAAAGCAGC ACAGATTCCA CATTTTTATA CATGAGGATC
TTCTTTGTGG TGAATACCAG GATTGACTGC ATCCCTTTAA AAGAAGTTTT
ATGTCCCTGA CTCTGGCTAA AATTATCTAA TTTCCAGATG CTTTTGTAGA
TGACTGAAGT ATTTGTGAGC CACATATTGG GAGTTCTAGA TTTGAGTGAA
TGGCAGGAAA GGGCCATCTC CATTGAGATG ATTAAGTGAA CCAAACTAGT
TCTCGGAATT CTACAGAGAA GGAGGGAATC AGACTGAGGA AGCTGTGACA
TAGGACTTGA AGACCAAAGA CTTTGAAATT TGCGAGCTGC TCATGTGTGA
GTTATTATCA CTGCTGTCTT TCTATTGAGT TACAAATCTA TATTTTTATT
GAAGTTTAAA TAAAGAAAAA ATTTACAAGA AAAAAAAAAA A

Fig 8

# GRO-1 and its human homologue hgro-1p

```
                                  .. .    . ...... .       ...  . .... . .... ..
hgro-1p          MASVAAARAVPVGSGLRGLQRTLPLVVILGATGTGKSTLALQLGQRLGGEIVSADSMQVYEGLDIITN
GRO-1                 MIFRKFLNFLKPYKMRTDPIIFVIGCTGTGKSDLGVAIAKKYGGEVISVDSMQFYKGLDIATN

                  .   .    ... ..  .    . .  ..    ..  . .. ...  .... .. .. .
hgro-1p          KVSAQEQRICRHHMISFVDPL-VTNYTVVDFRNRATALIEDIFARDKIPIVVGGTNYYIESLLWKVLVN
GRO-1            KITEEESEGIQHHMMSFLNPSESSSYNVHSFREVTLDLIKKIRARSKIPVIVGGTTYYAESVLYENNLI

                              .  ..       .  .  .  .  ...    . . .. . ....
hgro-1p          TKPQEMGTEKVIDRKVELEKEDGLV------LHKRLSQVDPEMAAKLHPHDKRKVARSLQVFEETGISH
GRO-1            ETNTSDDVDSKSRTSSESSSEDTEEGISNQELWDELKKIDEKSALLLHPNNRYRVQRALQIFRETGIRK
```

Fig 9A

```
                 ..   .  .   ... . .        .  .. .... ... .  ..  ..     .
hgro-1p          SEFLHRQHTEEGGGPLGGPLKFSNPCILWLHADQAVLDERLDKRVDDMLAAGLLEELRDFHRRYNQKNV
GRO-1            SELVEKQKSDETVD-LGGRLRFDNSLVIFMDATPEVLEERLDGRVDKMIKLGLKNELIEF---YNEHAE

                       .  .  . .. ...   .          . .
hgro-1p          SENSQDYQHGIFQSIGFKEFHEYLITEGKCTLETSNQLLKKGPGPIVPPVYGLE----------
GRO-1            YINHSKY--GVMQCIGLKEFVPWLNLDPSERDTLNGDKLFKQGCDDVKLHTRQYARRQRRWYRSRLLK

                   ..      .          .              .
hgro-1p          VSDVSKWEESVLEPALEIVQSFIQGHKPTATPIKMPYNEAENKRSYHL----------
GRO-1            RSDGDRKMASTKMLDTSDKYRIISDGMDIVDQWMNGIDLFEDISTDTNPILKGSDANILLN

                 .  .    .  .  ..  . .   .
hgro-1p          CDLCDRIIIGDREWAAHIKSKSHLNQLKKRRRLDSDAVNTIESQSVSPDYNKEPKGKGSPGQNDQELKCSV
GRO-1            CEICNISMTGKDNWQKHIDGKKHKHHAKQKKLAETRT
```

C2H2 zinc finger

Fig 9B

00S2220" TSTETS60

Conceptual translation of a partial sequence of the Drosophila homologue of *gro-1*

P I T C K H K K Q L T A T S G S V P I G I H V L K T C G F Y L P *Stop* L T *Stop* I H S Q *Stop* V E
**Met** I R K V P L I V V L G S T G T G K T K L S L Q L A E R F G G E I I S A D S **Met** Q V Y T H L
D I A T A K A T K E E Q S R A R H H L L D V A T P A E P F T V T H F R N A A L P I V E R L L
A K D T S P I V V G G T N Y Y I E S L L W D I L V D S D V K P D E G K H S G E H L K D A E L
N A L S T L E L H Q H L A K I D A G S A N R I H P N N R R K I I R A I E V Y Q S T G Q T



FIG 10

# Structure of pMQ8

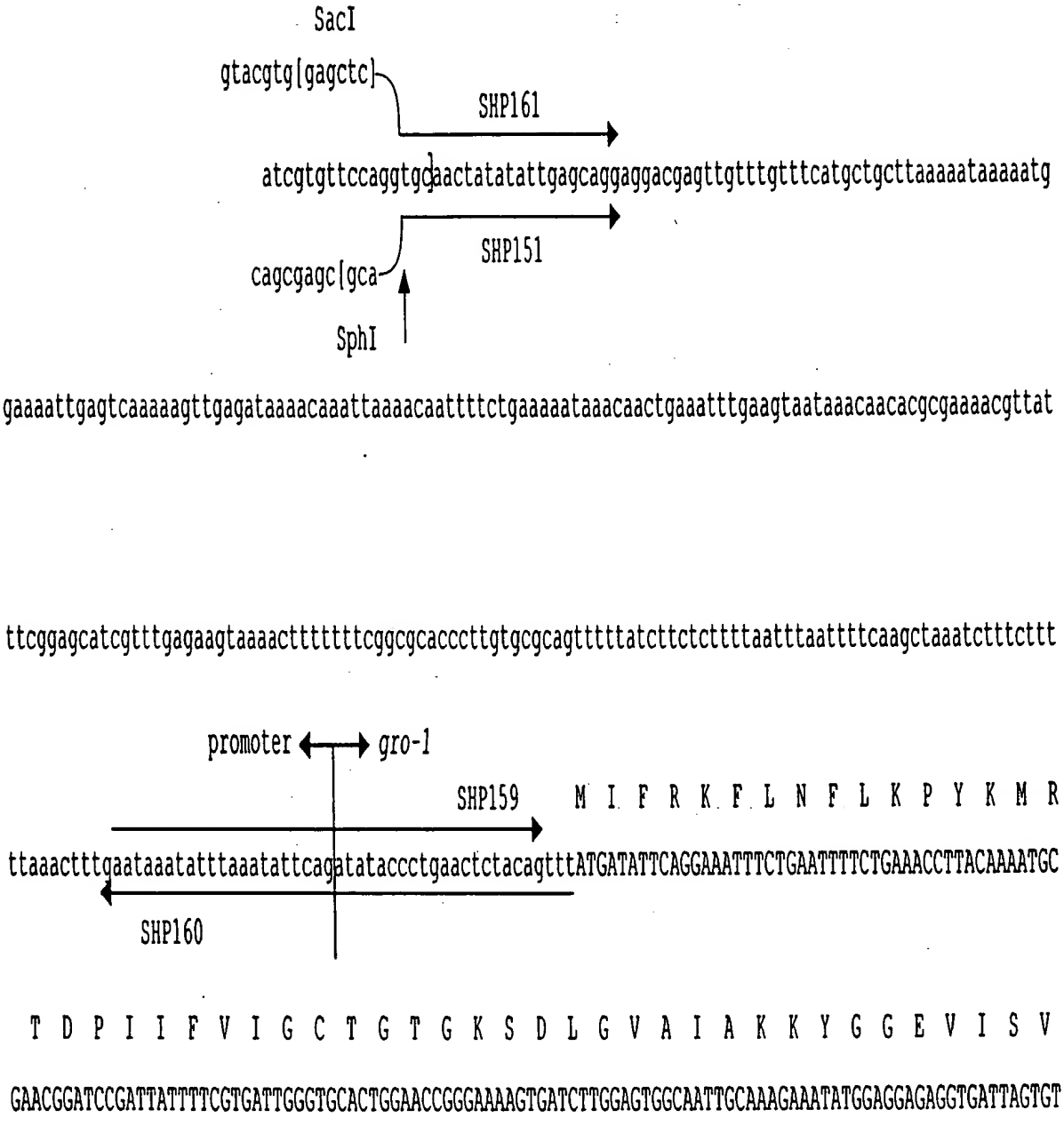



Fig. 11A

D S M Q F Y K G  L D I A T N . . .

AGATTCAATGCAATTTTATAAAGgtacatgggttttgtttcaattttaaattaattaattttcgtttttcagGACTTGACATTGCCACGAAT.....

. . . H A K Q K K L A E T R T •

.....CATGCTAAGCAAAAGAAATTGGCAGAGACTCGCACAtaagacgctatatttatttttgttaacttaaattatttttgttgttgattgtt

SHP170

[tctaga]tatact

XbaI

ctctaaataaaaaaacagctcagagagaagattaggcgctcgtccacatctccgacgatagtcaacccgaacgaagggaactatctttaattgtcagtga

SHP162

[ctgcag]tgtcat

PstI

Fig. 11B

# Construction of pMQ18

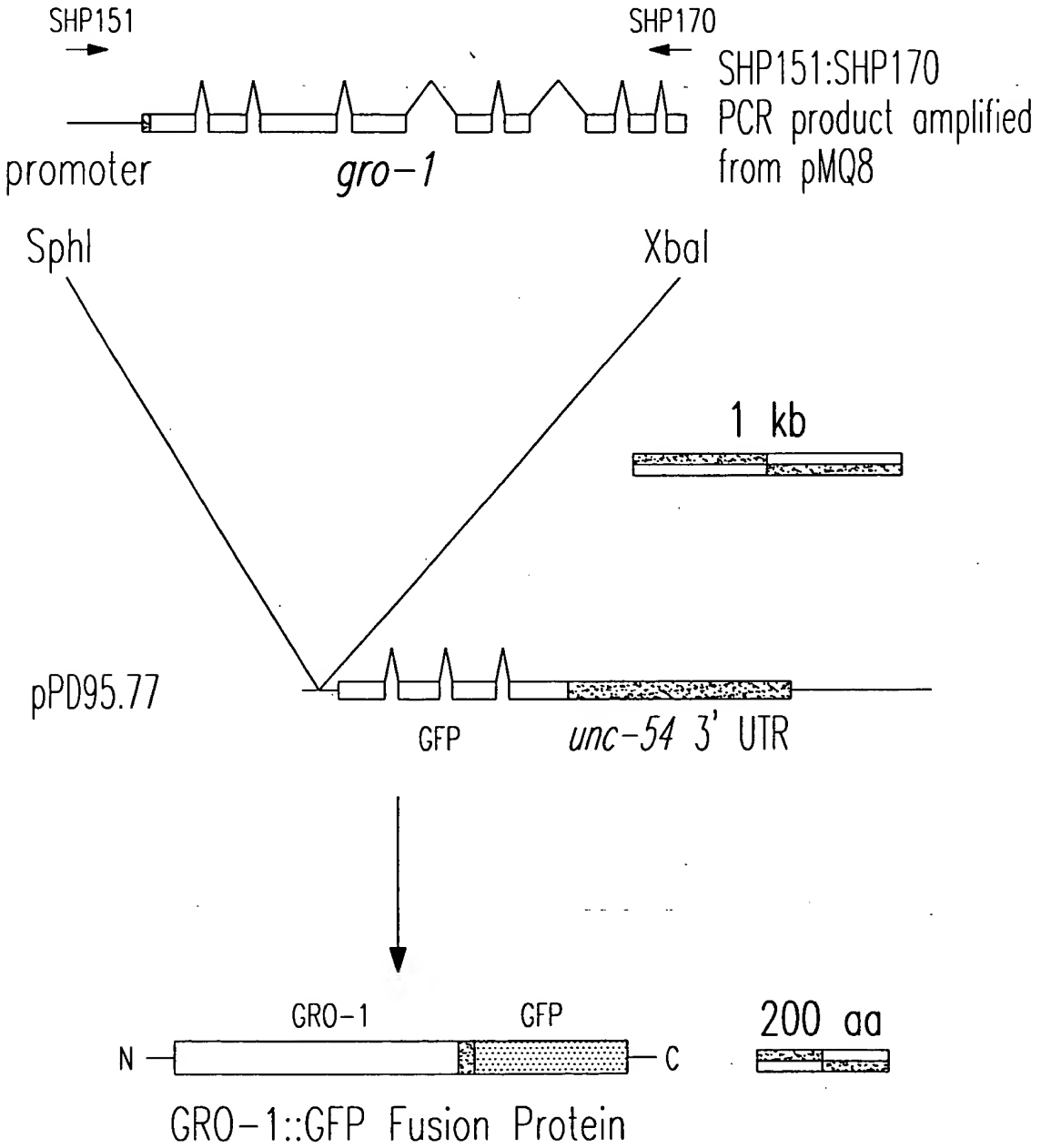



Fig. 12

gop-1

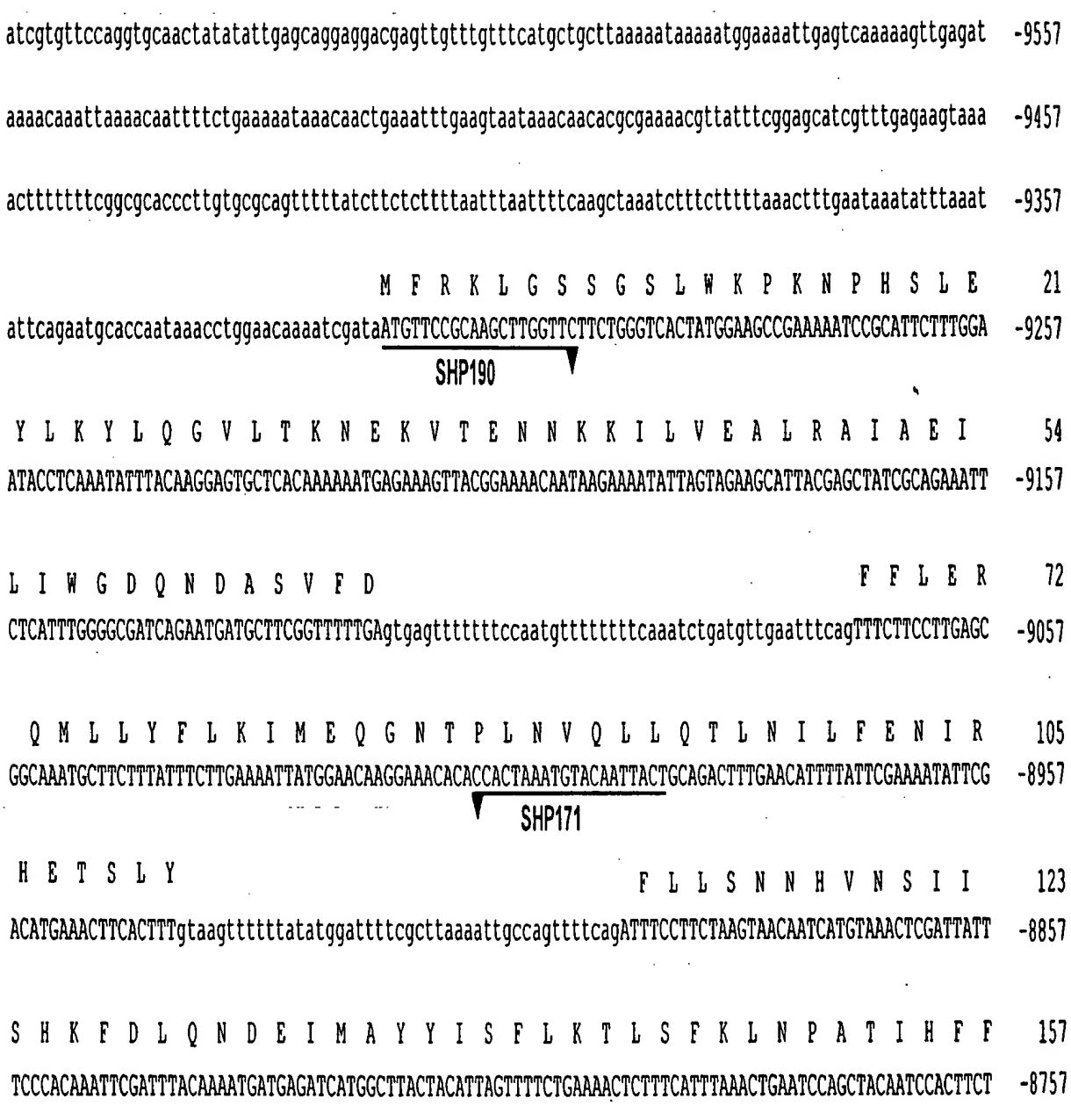

atcgtgttccaggtgcaactatatattgagcaggaggacgagttgtttgtttcatgctgcttaaaaataaaaatggaaaattgagtcaaaaagttgagat -9557
aaaacaaattaaaacaattttctgaaaaataaacaactgaaatttgaagtaataaacaacacgcgaaaacgttatttcggagcatcgtttgagaagtaaa -9457
actttttttcggcgcacccttgtgcgcagttttatcttctcttttaatttaattttcaagctaaatctttctttttaaactttgaataaatatttaaat -9357
M F R K L G S S G S L W K P K N P H S L E 21
attcagaatgcaccaataaacctggaacaaaatcgataATGTTCCGCAAGCTTGGTTCTTCTGGGTCACTATGGAAGCCGAAAAATCCGCATTCTTTGGA -9257
SHP190
Y L K Y L Q G V L T K N E K V T E N N K K I L V E A L R A I A E I 54
ATACCTCAAATATTTACAAGGAGTGCTCACAAAAAATGAGAAAGTTACGGAAAACAATAAGAAAATATTAGTAGAAGCATTACGAGCTATCGCAGAAATT -9157
L I W G D Q N D A S V F D F F L E R 72
CTCATTTGGGGCGATCAGAATGATGCTTCGGTTTTTGAgtgagttttttccaatgttttttttcaaatctgatgttgaatttcagTTTCTTCCTTGAGC -9057
Q M L L Y F L K I M E Q G N T P L N V Q L L Q T L N I L F E N I R 105
GGCAAATGCTTCTTTATTTCTTGAAAATTATGGAACAAGGAAACACACCACTAAATGTACAATTACTGCAGACTTTGAACATTTTATTCGAAAATATTCG -8957
SHP171
H E T S L Y F L L S N N H V N S I I 123
ACATGAAACTTCACTTTgtaagttttttatatggattttcgcttaaaattgccagttttcagATTTCCTTCTAAGTAACAATCATGTAAACTCGATTATT -8857
S H K F D L Q N D E I M A Y Y I S F L K T L S F K L N P A T I H F F 157
TCCCACAAATTCGATTTACAAAATGATGAGATCATGGCTTACTACATTAGTTTTCTGAAAACTCTTTCATTTAAACTGAATCCAGCTACAATCCACTTCT -8757

Fig 13A

*gop-1* continued...

```
 F  N  E  T  T  E  E  F  P  L  L  V  E  V  L  K  L  Y  N  W  N  E  S  M  V  R  I  A  V  R  N  I  L   190
TCTTCAATGAAACGACTGAAGAATTTCCATTGTTGGTAGAAGTTTTGAAGCTTTATAATTGGAATGAATCAATGGTTCGAATTGCTGTTAGAAATATTCT  -8657
                                  SHP141                                SHP172
 L  N  I  V  R  V  Q  D  D  S  M  I  I  F  A  I  K  H  T  K                                          210
TTTAAATATTGTGAGAGTTCAAGATGATTCAATGATTATTTTCGCTATCAAGCATACAAAAgttagtagaaaattattttgaaaaggtgtatttaagcaa  -8557

                    E  Y  L  S  E  L  I  D  S  L  V  G  L  S  L  E  M  D  T  F  V  R  S  A  E  N  V  L  A  N   240
taaatattacagGAATATCTATCGGAGTTAATAGATTCTCTAGTTGGTCTCTCACTTGAAATGGACACATTTGTACGATCTGCTGAGAATGTGTTAGCTA  -8457
                                                                                                    ▼

 R  E  R  L  R  G  K  V  D  D  L  I  D  L  I  H  Y  I  G  E  L  L  D  V  E  A  V  A  E  S  L  S  I   273
ATCGAGAGAGATTACGAGGAAAAGTGGATGATTTAATTGATTTGATTCATTATATTGGTGAACTATTGGATGTGGAAGCTGTCGCCGAAAGTTTATCAAT  -8357
SHP142                                SHP173
 L  V                                                T  T  R  Y  L  S  P  L  L  L  S  S  I  S  P  R   291
TTTAGgtcagttttactgctggaaaatcaagtttttaatgttaaattttcagTAACAACACGATACTTAAGCCCTCTATTACTTTCAAGTATATCACCAA  -8257

 R  D  N  H  S  L  L  L  T  P  I  S  A  L  F  F  F  S  E  F  L  L                                    313
GAAGAGATAATCATTCACTTCTACTCACTCCGATTTCTGCGTTATTTTTTTTCTCTGAATTTTTATTGgtgagttttaacatttaaaattacatttttct  -8157

                    I  V  R  H  H  E  T  I  Y  T  F  L  S  S  F  L  F  D  T  Q  N  T  L  T  T  H  W  I   341
aatttatttatttttcagATAGTTCGTCACCATGAAACAATATATACATTTTTATCATCTTTCCTATTTGACACTCAGAATACTTTGACGACCCATTGGA  -8057

 R  H  N  E  K  Y  C  L  E  P  I  T  L  S  S  P  T  G  E  Y  V  N  E  D  H                            366
TACGTCATAATGAGAAATATTGCTTAGAACCGATTACATTATCATCACCAACCGGAGAATATGTGAATGAAGACCAgtaagagctgaaattttaaaattt  -7957

                          V  F  F  D  F  L  L  E  A  F  D  S  S  Q  A  D  D  S  K  A  F  Y  G  L  M   391
ttgctttgaatatagtattttcagCGTATTTTTCGATTTTCTACTGGAAGCATTTGATTCCAGTCAAGCAGACGATTCGAAGGCATTCTATGGATTAATG  -7857
```

Fig 13B

# *gop-1* continued...

```
L  I  Y  S  M  F  Q  N  N  A                                                    401
CTGATTTATTCAATGTTTCAGAATAATGgtgagttttaaaaaattgatttgttaaattaaaatttccatttccaataactcctcttcagacagtaagttt -7757

tcaatgttgtaaagttcctgttcatctgtgatcgttttcttcatttttttagttttgcatgaacagttttcaaatttttttgatatcatacagtaaatat -7657

cgtcatccagataattttctatttaaaaaaaatgaataaaaagagggcgcgcagaaattgccgaagtaatgtaaatttaaagggacacatgcgtagcttg -7557

ttgtgtgggtctcgccgcgctttgtttgatttatcttgttttctgctcaaagagctgtttttatttttagcgttgaatgctttttttaccgttctcatcggc -7457

tttttaataggaatatttaaaaaaaaaggtttaataaatcttcgtttttacaaaatccatctaagatttgcatttgtgaagctcaacaagtaaagtttta -7357

agtaacattgtttttttaaaaaacaattgaaccaaattttgccgaaacattaataacatgacgatactctataaaatattcctcttttcaaaataaatttt -7257

                                                  D  V  G  E  L  L  S  A  A  N  F  P  V  L  K  E  S  T  T  T  S  L  A  Q  Q  N   427
caaaaaaaatccatttttcagCCGATGTTGGAGAACTTCTATCTGCTGCCAACTTCCCAGTGCTCAAAGAATCAACGACAACTTCATTAGCTCAACAGAA -7157
                                                                  SHP174

L  A  R  L  R  I  A  S  T  S  S  I  S  K  R  T  R  A  I  T  E  I  G  V  E  A  T  E  E  D  E  I  F   480
TCTTGCTCGTCTCCGAATAGCATCTACGTCTTCCATATCAAAGCGAACGAGAGCTATCACTGAAATTGGAGTAGAAGCGACCGAGGAAGATGAGATTTTT -7057
                                                                                  SHP185

H  D  V  P  E  E  Q  T  L                                                       469
CATGATGTTCCTGAAGAACAAACGTTGgtaagtaaataaatcaacattgattgttacacaaactttaatatttttaaatttgaaaattttcttcaaagtg -6957

                                                E  D  L  V  D  D  V  L  V  D  T  E  N  S  A  I  S  D  P  E   489
ctcaaaaatcctgtcgaaaattacagGAAGATCTGGTGGATGATGTATTGGTTGATACTGAAAATTCAGCAATAAGTGATCCAGAAgtgagtagaaaacg -6857

                                                                                 P  K  N  V  E  S  E  S  R   498
tgcatgtattaattattaaaaaaaaaatatagttttccccagttttccttgacctaaaactcagcaatttcagCCTAAAAACGTGGAGTCAGAATCTCGT -6757
```

FIG. 13C

*gop-1* continued...

```
S R F Q S A V D E L P P P S T S G C D G R L F D A L S S I I K A V G  532
TCTCGATTTCAATCTGCTGTTGATGAGCTTCCACCTCCGTCGACTTCTGGATGTGATGGTCGACTTTTTGATGCACTTTCATCGATTATCAAAGCAGTTG -6657

 T D D N R I R P I T L E L A C L V I R Q I L M T V D D E K          561
GAACAGATGACAATCGAATTCGACCAATTACATTGGAACTTGCATGTCTTGTAATTCGGCAAATTTTAATGACTGTTGATGATGAAAAAgtaagattaca -6557
          SHP175
                                                V H T S L T K L C F E V R L K L L S  579
aattcaaaattgagcaaaatcagaatctaaatttcataaattgttcagGTACATACCAGTTTAACGAAATTATGCTTCGAAGTTCGTCTAAAACTTTTAT -6457

 S I G Q Y V N G E N L F L E W F E D E Y A E F E                    603
CATCAATTGGACAATATGTTAATGGAGAGAATCTGTTTTTGGAGTGGTTTGAGGATGAATATGCAGAATTTGAAgtaagccaagaggtccgaaaataatt -6357

                    V N H V N F D I I G H E M L L P P A A T P L S N L L L  630
taattcatccttttattcagGTGAATCACGTGAATTTCGATATAATCGGTCACGAAATGCTTCTTCCTCCAGCTGCAACTCCTCTTTCGAATCTGCTAC -6257

 H K R L P S G F E E R I R T                                Q I V  647
TTCATAAGCGATTGCCCAGTGGATTTGAAGAACGAATAAGAACTgtaggaaacttttaaatttgaaaattaattatatatatatttgcagCAAATCGTA -6157

F Y L H I R K L E R D L T G E G D T E L P V R V L N S D Q E P V A I  681
TTCTACCTACATATTCGAAAATTGGAACGAGATTTGACCGGTGAAGGAGACACAGAATTACCTGTGAGAGTGTTGAATTCTGATCAGGAACCAGTTGCCA -6057

 G D C I N L H                                        N S D L L S C T  696
TCGGTGATTGTATTAATTTACgtgagttcatctgcatagaaaacaccatatttctactcaaattaacaattttcagATAATTCGGATCTTCTATCCTGCA -5957

 V V P Q Q L C S L G K P G D R L A R F L V T D R L Q L I L V E P D  729
CTGTGGTTCCTCAACAACTATGTTCTCTTGGAAAACCTGGTGATCGTCTTGCTCGATTCCTTGTCACTGATAGACTTCAATTAATTCTTGTCGAACCGGA -5857
          SHP176
 S R K A G W A I V R F V G L L Q D T T I N G D S T D S K V L H V V  762
TTCTCGAAAAGCCGGATGGGCAATTGTTCGATTCGTAGGACTTCTTCAAGATACAACAATTAATGGAGATTCTACGGATTCGAAAGTTTTGCATGTTGTG -5757
                                                    SHP177
 V E G Q P S R I K                                        K R H P V L T A  779
GTGGAAGGGCAACCCTCGAGAATTAAGgtaagaatactaacgggaaaaaaaaatcaaaaaattacttctgtttcagAAAAGACATCCGGTTTTAACTGCA -5657
```

FIG 130

# *gop-1* continued...

```
A F I F D D H I R C M A A K Q R L T K                                    798
AAGTTCATATTCGATGATCACATTCGGTGTATGGCAGCAAAGCAACGGCTCACCAAGgtaacggaaaaaataaccaaaaagacggaaagttattgtaaat -5557

ggacgaaatcggcgaaattaattgaaaacgtttgaatttgccgctaaaaccaaacgaaaaccaaacgaaagcgaaatttaactatcccttcaggtagaat -5457

                    G R Q T A R G L K L Q A I C S A L G V P R I D P A T   824
atacattttatttctctttatagGGTCGCCAAACAGCACGTGGTCTGAAACTTCAGGCGATATGTTCAGCTCTTGGAGTTCCACGTATCGATCCAGCGAC -5357

 M T S S P R M N P F R I V K G C A P G S V R K T V S T S S S S S Q    857
AATGACGTCATCACCACGAATGAATCCATTCAGAATTGTGAAAGGATGCGCACCGGGAAGTGTACGAAAAACTGTTTCCACATCATCATCGTCAAGCCAA -5257

G R P G H Y S A N L R S A S R N A G M I P D D P T Q P S S S S E R R   891
GGACGTCCCGGACATTATTCTGCAAATCTTAGATCAGCATCTAGAAATGCAGGAATGATACCAGATGATCCAACTCAACCGAGTAGTTCTTCGGAAAGAA -5157
                                  SHP178
 S •                                                                    892
GATCCtagggatcaatatctcttcagtttcatcattttatgctgtaaattgtatttaagtattcctattctttgtagtactgtatttacacatcgtctag -5057

ttaaaatcacaaatctccgaaaaaacaaaccagtgaacatgtgatatttctcttgcccatagttctctttttttttttgaaacaaaaacaattacttttat -4957
                                                                  polyA
gctcacctattcgagccatatttttttcccaattaccggttgtttatttaatttctttttttttttctgtaaatctactttatttttaaaactgcatttg -4857

agattgtgtatatttttttcaaaatggttcaaatgccgaatctatctactt                     -4807
```

FIG 13E

gop-2

SL2

M A E K A E N L P S S S A E A S E 1

tttaatcattattcaaacagaaaaaccgattatttattcagattctcaaaaATGGCTGAAAAAGCTGAAAATCTTCCATCTTCTTCGGCCGAAGCTTCAG -470

E P S P Q T G P N V N Q K P S I L V L G M A G S G K T T F V Q 4

AAGAGCCATCACCTCAAACTGGACCAAATGTGAATCAAAAACCATCGATTTTGGTTCTTGGAATGGCTGGTTCTGGAAAAACGACATTTGTTCAGgtaac -460

R L T A F L H A R K T P P Y V I N L D P 6

tttcattcaattttgagagttttcaaacattactattttcagCGTCTCACAGCATTCCTACATGCTCGTAAAACACCTCCATATGTGATTAATCTGGATC -450

A V S K V P Y P V N V D I R D T V K Y K E V M K E F G M G P N G A 10

CGGCAGTTAGCAAAGTACCTTATCCAGTGAATGTTGACATTCGAGATACTGTGAAATACAAGGAAGTTATGAAAGAATTCGGAATGGGACCAAATGGAGC -440

SHP179

I M T C L N L M C T R F D K V I E L I N K R S S D F S V C L L D T 13

AATTATGACATGTCTTAACCTGATGTGTACTCGTTTTGATAAAGTAATTGAGTTGATTAATAAGAGATCTTCTGATTTCTCAGTTTGTCTTCTTGATACT -430

SHP180

P G Q I E A F T W S A S G S I I T D S L A S S H P T 16

CCTGGACAAATTGAAGCATTCACTTGGAGTGCTAGTGGATCTATTATCACTGATTCATTGGCAAGTAGCCATCCCACGgtaagggattttgatttatgaa -420

SHP143

atctgcttgaaatgaaaaaagattctaataaatttttgacttttaaacatttttttacagttatatttggtctattttctatcattaaaagcaaaatgaaa -410

V V M Y I V D S A R A T N P T T F M S N 18

agtcgattctactccatatttattaatttcgacttttcagGTGGTAATGTACATTGTGGATTCCGCTCGTGCCACAAATCCAACTACATTCATGTCCAAT -400

SHP144

Fig 14A

## *gop-2* continued...

```
M L Y A C S I L Y R T K L P F I V V F N K A D I V K P T F A L K W M  21
ATGCTCTACGCATGTTCCATTCTCTACCGTACCAAACTTCCATTCATTGTCGTTTTCAACAAAGCTGATATTGTCAAACCAACATTTGCACTCAAATGGA -390

 Q D F E R F D E A L E D A R S S Y M N D L S R S L S L V L D E F Y  24
TGCAAGATTTCGAAAGATTTGATGAAGCTTTAGAGGATGCCAGAAGCAGTTATATGAATGATTTGAGTCGTTCATTGAGTCTCGTTCTTGATGAATTCTA -380
                  ▼ SHP181

 C G L K T V                                     C V S S A T G E G F E D V  26
TTGCGGACTGAAAACAGgtttttattcgaaataaaacctttttaaataataaatttcagTTTGCGTCAGTTCTGCAACTGGAGAAGGATTCGAAGATGT -370

 M T A I D E S V E A Y K K E Y V P M Y E K V L A E K K L L D E E E  29
AATGACAGCAATCGATGAAAGTGTTGAAGCATACAAAAAAGAATATGTTCCAATGTATGAAAAAGTGTTGGCTGAGAAAAAACTATTGGATGAGGAGGAG -360

R K K R D E E                                   T L K G K A V H D L N K V  31
AGAAAGAAAAGAGATGAAGAGgtaattgtagtaatttaattctgattatcttcaaattttcagACTCTGAAAGGAAAAGCTGTTCACGACCTGAACAAAG -350

 A N P D E F L E S E L N S K I D R I H L G G V D E E N E E D A E L  35
TCGCCAATCCCGACGAATTTCTGGAGTCGGAGTTGAATTCAAAAATCGATAGAATTCATTTGGGCGGAGTCGATGAAGAGAATGAGGAGGATGCTGAACT -340
          ▼ SHP182

 E R S •                                                          35
CGAAAGATCCtgattttcttttgtttttgaatttttattctattttgatccctgtttacttcttattgttctcattttgttgcgttgttttacatttta -330

                    polyA
                    ⌐
ctcatttttgcataaacttgttgcaaaaatcaatataatttttgatctggaaatggttttaaaccttaacctttcatatattaataatttttttcaaaa -320

aaacgttctaaaaaggttcctcatttttttcaatataggaaattttgaaga                                              -315
```

FIG - 14B

*gop-3*

SL2

```
                                                                M  S  E  K  T  F  H  K    8
tcttttccaaaaatgaggttcttcgcttgaaaagccaacatttaaaacctttttttttccagaaacctagtggttaATGTCTGAAAAGACGTTCCACAAG  -3057

A  Q  T  I  R  A  K  A  S  G  V  P  S  I  V  E  A  V  Q  F  H  G  V  R  I  T  K  N  D  A  L  V  K  E   42
GCACAGACCATCCGTGCAAAGGCATCCGGAGTGCCTTCAATCGTCGAAGCTGTACAGTTTCATGGAGTTCGCATCACAAAAAACGATGCTTTGGTTAAGG  -2957

                                                                                    V  S  E  L  Y  R   48
AGgtactacccaaatttcaaaatgttgcacaattcaattgaaaatataaattgtgaattaaattcaacttacatgttttttcagGTTTCCGAATTATACA  -285

 S  K  N  L  D  E  L  V  H  N  S  H  L  A  A  R  H  L  Q  E  V  G  L  M  D  N  A  V  A  L  I  D  T   81
GAAGTAAAAATCTAGATGAACTTGTTCATAACTCTCATCTGGCGGCTCGTCATCTTCAAGAAGTTGGATTAATGGATAATGCAGTTGCTCTAATTGATAC  -275
                                                                SHP183

 S  P  S  S  N  E  G  Y  V  V  N  F  L  V  R  E  P  K  S  F  T  A  G  V  K  A  G  V  S  T  N  G  D   114
ATCTCCAAGCTCAAATGAAGGATATGTTGTCAATTTCCTAGTTCGAGAACCAAAATCATTCACTGCTGGAGTCAAAGCAGGAGTTTCAACGAATGGAGAT  -26

A  D  V  S  L  N  A  G  K  Q  S  V  G  G  R  G  E  A  I  N  T  Q  Y  T  Y  T  V  K                    14
GCGGATGTCAGTTTAAATGCCGGAAAACAAAGTGTTGGAGGACGAGGAGAGGCAATCAATACACAGTATACATATACTGTAAAGgtaaggacgagagttg  -255
                                SHP145

gcactgccagtttggcatgttctcccaatatttttaattataaaatttggaagtataaaaaaatgtttgcttcatctaaaaatagccttttcacatga  -245

aaaaaattgaaaaaaagtgctcaaaaatttcagaaatttccaatttccaaacaattttggagaactttcaaaaatttttccaactgaaattaaagctata  -235
```

FIG. 15A

*gop-3* continued...

```
                                                                G  D  H  C  F    147
ttctatcactaaattttatacaagtcttaagagaaaatgatgaagtggctcattttgtagaatttcctaaaaaataatatcttcagGGCGATCACTGCTT   -225

 N  I  S  A  I  K  P  F  L  G  W  Q  K  Y  S  N  V  S  A  T  L  Y  R  S  L  A  H  M  P  W  N  Q  S    180
CAACATTTCCGCAATCAAACCATTCCTGGGATGGCAAAAATATTCGAATGTATCAGCGACTCTATACCGTTCACTTGCACATATGCCATGGAATCAATCA   -215
     SHP138                                                                                  SHP146

 D  V  D  E  N  A  A  V  L  A  Y  N  G  Q  L  W  N  Q  K  L  L  H  Q  V  K  L  N  A                    208
GATGTTGATGAGAATGCAGCTGTTCTTGCATATAATGGACAACTATGGAATCAAAAGCTTTTGCATCAAGTCAAATTGAATGCGgtaaagtattataagt   -205

                                                 I  W  R  T  L  R  A  T  R  D  A  A  F  S  V  R  E  Q  A  G  H  T  L     23
gttttgtccaaactatgatacagttcttcagATATGGAGAACACTTCGTGCCACTCGAGATGCCGCATTTTCAGTTCGTGAACAAGCCGGACACACTTTG   -195

 K  F  S  L  E  N  A  V  A  V  D  T  R  D  R  P  I  L  A  S  R  G  I  L  A                             25
AAATTCTCGTTGGAGAATGCTGTAGCTGTTGATACAAGAGATAGACCTATTCTTGCAAGTCGTGGAATTCTTGgtaagagtaacaacgactatttttaaa   -185

aaatatctttttcgaaaaaattacgaacgaaaaaaaactgtattatgtacccaaacgcgaaattttgcagttcttgcgcgttcttgttgataaaaaatat   -175

                                                                R  F  A  Q    26
gtaaaaaattggaaaaactacgaaaagtcgataaaaattccgtaccaaccggaaaatgtttcattaatttctcttccttttttcagCTCGTTTTGCTCAA   -165

 E  Y  A  G  V  F  G  D  A  S  F  V  K  N  T  L  D  L  Q                                              279
GAGTACGCAGGAGTATTTGGTGATGCGTCATTTGTGAAGAATACATTAGATTTACAGgtaacaaccttatttcaacaattatttcaaattctattaaaaa   -155
          SHP139

       A  A  A  P  L  P  L  G  F  I  L  A  A  S  F  Q  A  K  H  L  K  G  L  G  D  R  E  V  H  I  L     31
taattccagGCAGCTGCCCCTCTTCCACTCGGTTTCATTCTTGCCGCCTCATTCCAAGCGAAACATTTGAAAGGACTCGGAGATCGAGAAGTTCATATTT   -145
          SHP140
```

Fig 15B

# *gop-3* continued...

```
D R C Y L G G Q Q D V R G F G L N T I G                                  330
TGGATAGATGTTATTTGGGTGGACAACAGGATGTTCGAGGATTTGGTCTGAATACTATTGGAgtgagttttaacgaaattctcttgaaagtcaaataatc -1357
SHP184

V K A D N S C L G G G A S L A G V V H L Y R P L I P P N M L F  361
attttcagGTTAAAGCAGATAACAGTTGTCTTGGAGGAGGTGCTTCACTTGCTGGTGTCGTTCATTTGTATCGGCCATTGATTCCACCAAATATGCTATT -1257

A H A F L A S G S V A S V H S K N L V Q Q L Q D T Q R V S A G F G  394
TGCACACGCATTCCTTGCATCTGGAAGTGTTGCATCAGTTCATTCCAAAAATTTGGTGCAACAATTACAGGATACTCAACGAGTATCAGCCGGATTTGgt -1157
SHP163

gagtttgaaatttaggaaacatttggatgaaatgtatttttaaaaatagatcagctttatttatttgaaaaaaaacgctcattaatcaatagtgatagt -1057

tccattctgagtttcttcttcttcctcgcggaatacaatttttgacttgttcgcatccttcttgtgtactttgtcaccaatcttctcatcaactaaatct -957

cgaaactgaaaaaatttcaaaattattccaaaaaatattgatgcagactacctttttgatggcttctggtacgtttctagcgtcgaatggattggctcct -857

ccaataattaaagtctcgttcggtagtttagccagacggacggtgtgcttcaacatttttctaattaatctatttcaattcaagtcactcactctctctt -757
```

Fig 15C

# *gop-3* continued...

```
gacgtcttcttctatattccaagaactctgcagaaaatccgtgtccgccttgtgtgtttctagttggcgtcggaggattcacgggtccaagacgaatgga  -657

tgtctaaaaaatgttatatttttgcataaagaaaacaccataccttcaccacttttgagttgtgggcgttctgaatggaattgatcgattattattgct  -557

ctttcttgatttgcttctatcagctgcgtaatgaggtgttctaaagatcagctttaattcatttggacaagtgctcctctaataaacttaccctgtactc  -457

atttttgaaacgatttacgatgataagattgaaagtggaagttaaatttagtctttcaaagttgaaataaaatcttcataaataaataaatttaaatgaa  -357

                                                                                    L  A  F  V  F  K  S  401
agattaaataaattaacgttcacgtagttaaaaaaataatttaaatcttaaacttctaataaaaaatctcaattttccagGACTCGCATTCGTGTTCAAAA  -257

 I  F  R  L  E  L  N  Y  T  Y  P  L  K  Y  V  L  G  D  S  L  L  G  G  F  H  I  G  A  G  V  N  F  L   434
GTATTTTCCGGCTGGAACTCAACTACACGTATCCATTGAAATATGTGCTCGGCGATTCATTGCTCGGTGGATTCCATATTGGAGCTGGTGTCAACTTCTT  -157

Gtagagattaattggatgcaagcacccctcaaaaagatttttttgaaaaacgataaattcacagaatttcagttctttttctccccctttattgttatt  -57
          SHP134

ttcatcgtaatgctgtgctagaagtcagagtaaatatgagtttttttgtgttctaggaattccatttttttcaggaagcaaatttaataaaaattatcgaa  44
          SHP164
     polyA
tttcttgctctaaagatgttgtacattttatggaaatgttcgtatagtaa                                                  94
                                SHP135
```

FIG. 15D

*hap-1*

SL2

M S L R K I N F V T G 11

ttcgaacactttatatttctcgttttaaaactgtcggtgttttatagtaaactatcttcagaaaaaaATGAGCCTACGAAAAATCAATTTCGTAACTGGA 194

SHP91 SHP118

N V K K L E E V K A I L K N F E 27

AACGTGAAGAAGCTTGAAGAAGTCAAGGCTATTTTGAAGAATTTCGAGgtaaaatatatttgatattattcgaacgcgaaattttgcgccaaaagtacga 294

tgcctggtctcaacacgacaatattttgttaaatacaaacgaatgtgcgccttcaaagaaaagtttcaatctttcgttgccgtggagatattttagagt 394

V S N V D V D L D E F 38

ttttgtttaaattatatatttgtcgtatcgaaaccgggtaccgtaatcaatcaattaaatattttcagGTTTCAAACGTGGATGTCGATTTGGATGAATT 494

SHP165

Q G E P E F I A E R K C R E A V E A V K G P V L 62

CCAAGGAGAACCCGAATTTATTGCCGAAAGAAAGTGCCGTGAGGCTGTTGAAGCTGTAAAAGGGCCCGTTTTGgtatggaaaattgtatttgttctaaaa 594

V E D T S L C F N A M G G L P G P Y I K W F L K N L K P E 91

attgtcaaatttcagGTCGAAGACACAAGTTTATGCTTCAACGCAATGGGCGGTCTTCCTGGACCTTATATCAAGTGGTTTTTGAAGAATTTGAAACCAG 694

SHP129

Fig. 16A

# *hap-1* continued...

```
  G  L  H  N  M  L  A                                                G  F  S  D  K  T  A  Y  A  Q  C  I  F   111
AAGGACTACATAATATGCTAGgtaaatattttaatttttgaaaaaacttatttttcagCCGGATTTTCTGACAAAACCGCCTATGCTCAATGCATCTTT  794

  A  Y  T  E  G  L  G  K  P  I  H  V  F  A  G                                                          126
GCGTACACTGAAGGACTCGGAAAACCTATTCATGTATTTGCTGgtatgatttttgaatttaattctttaattttatatgttaatttagttgtttcattc  894

                                                   K  C  P  G  Q  I  V  A  P  R  G  D  T  A  F  G  W  D  P   145
ctcaatttatgagagatttttttttcaatttttctatttcagGAAAATGTCCTGGTCAAATTGTTGCTCCACGTGGTGATACTGCTTTTGGATGGGATCC  994
                                                                                  ▼ SHP130
  C  F  Q  P  D  G  F  K  E  T  F  G  E  M  D  K  D  V  K  N  E  I  S  H  R  A  K  A  L  E  L  L  K   178
ATGCTTCCAGCCAGATGGTTTTAAAGAAACATTCGGAGAAATGGATAAAGATGTAAAAAATGAAATTTCTCATCGTGCAAAGGCTCTGGAACTCCTCAAG  1094
                  ▼ SHP119                                                    ▼ SHP120

  E  Y  F  Q  N  N  •                                                                                    184
GAATATTTTCAGAATAATtaaattatttttttctcatctatgcaatttcttgaaaatttgttaagtttccgttgttatgcatttgcttttatttaaaaaaa  1194

          polyA
          ⌜
aaagaatatttttacattaatattagatatgagaaaagagtaatttctggattttaaccttcctacaaaagaatatttatatttttttgtatgatttttta  1294
                                                  ▼ SHP93
```

Fig. 16B